**EXPERIMENT 4:**

**Implement the Word Count program using Hadoop MapReduce**

Zoya Momin

Department of Computer Engineering

M.H Saboo Siddik College of Engineering

Mumbai, India

[zoya.221257.co@mhssce.ac.in](mailto:zoya.221257.co@mhssce.ac.in)

1. **INTRODUCTION**

**MapReduce** is the processing engine of Hadoop. While HDFS is responsible for storing massive amounts of data, MapReduce handles the actual computation and analysis. It provides a simple yet powerful programming model that allows developers to process large datasets in a distributed and parallel manner. This approach is highly scalable and can be executed on frameworks like Hadoop, making it suitable for processing massive text datasets efficiently. It demonstrates how distributed computing can solve simple yet computationally intensive tasks effectively

**Start the Cloudera**

Graphical user interface, text, application

Description automatically generated

**Open Eclipse**

Graphical user interface, text

Description automatically generated

**Create a New Project:**

Graphical user interface, text

Description automatically generated

**Open Add external JARs**

Graphical user interface

Description automatically generated

Graphical user interface

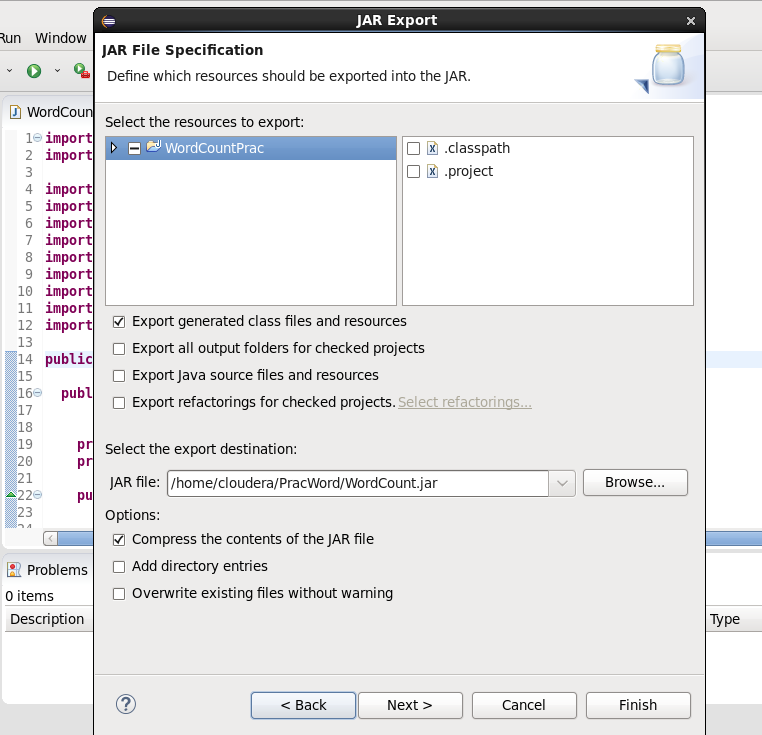
Description automatically generated

**Now click on Finish button**

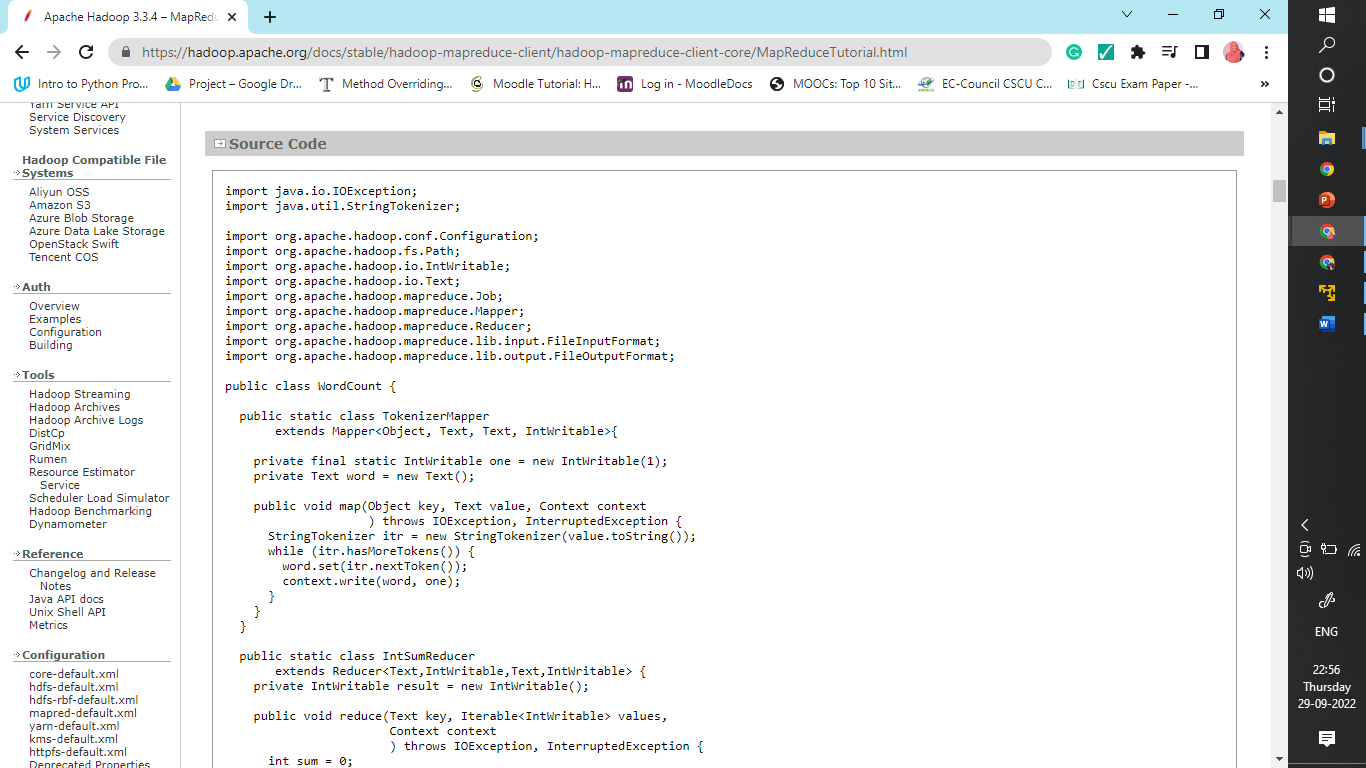
Graphical user interface, text, application

Description automatically generated

**Open file**

****

**Create Class with Class Name as WordCountProgram and click Finish Button**



Graphical user interface, text, application

Description automatically generated

**Step 7: Open the Terminal**Graphical user interface, text, application

Description automatically generated

**Fig 19: Terminal**

**1. The working directory:**

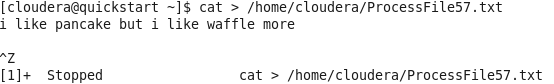
image1.png

**2. List the content of /home/cloudera**

Text

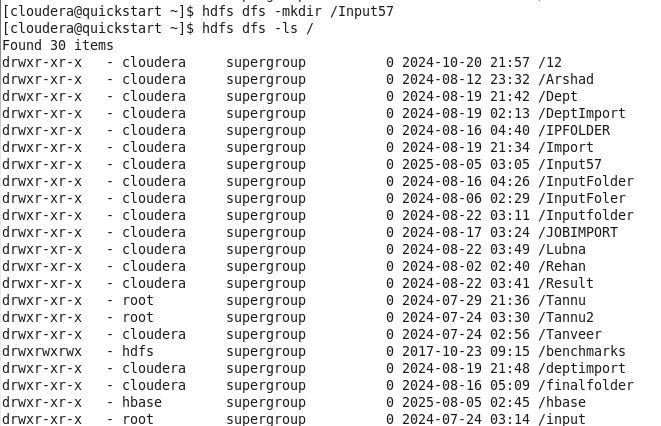
Description automatically generated with medium confidence

3**. Create a file ProcessFile57.txt and Display the of the ProcessFile57.txt using cat command**

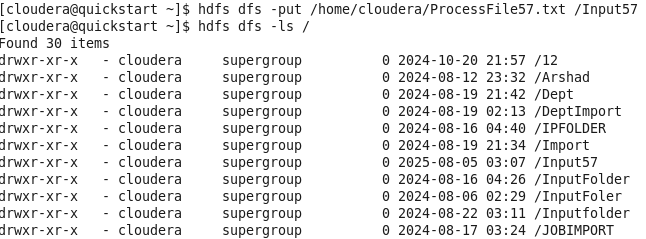
****

**4. Check the content of hadoop file system:**

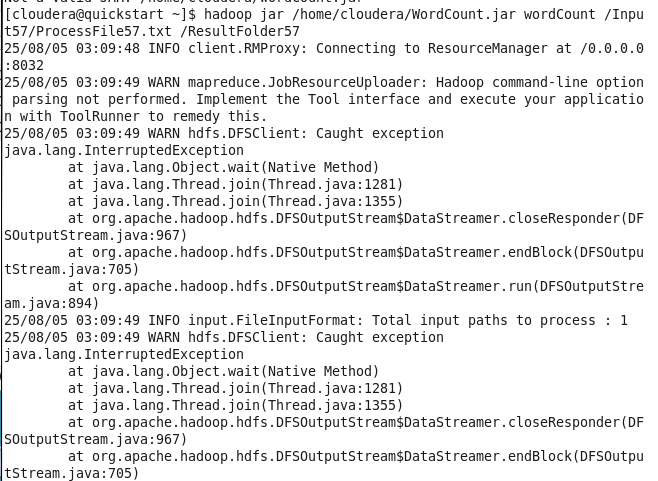
You can use either of the command



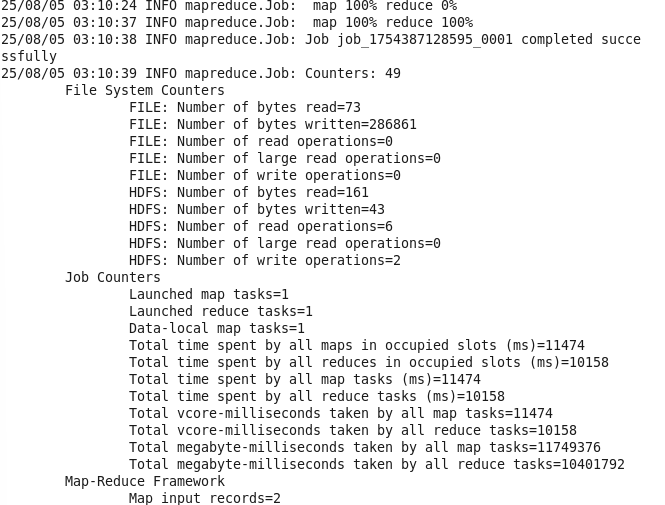
**5. Create a folder in Hadoop file system**

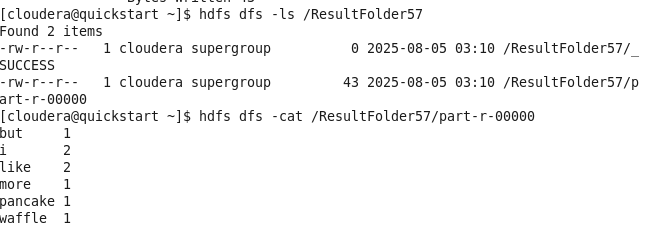
****

6**. Run the WordCount.jar File on hadoop**



7**. View the Map-Reduce Output**

****

****